CLAIMS:

1. A Link Packet Scheduler for each virtual lane (VL) at a given port, comprising:

a N-bit counter arranged to accumulate free credits relinquished, when a data packet is removed from a receive buffer and that buffer space is reclaimed as available for data packet storage, or when a link packet is received whose Flow Control Total Bytes Sent (FCTBS) field differs from actual blocks received (ABR) at the given port;

a first comparator arranged to make comparison between accumulated free credits from the N-bit counter and a programmable credit threshold;

a second comparator arranged to make comparison between a current buffer receive utilization indicating a data storage level of the receive buffer and a programmable utilization threshold; and

a logic device arranged to track a current link state of the corresponding port, to monitor the amount of receive buffer resources from the first and second comparators and to schedule for the transmission of a link packet, via a physical link.

2. The Link Packet Scheduler as claimed in claim 1, wherein said logic device comprises:

a first OR gate arranged to track whether a Link State Machine transitions into one of the ***LinkInitialize***, the ***LinkArm*** and the ***LinkActive*** state, whether a configuration strap for enabling

loopback operation changes states, and whether a per-VL Link Packet Watchdog Timer expires, and generate therefrom a logic signal;

an AND gate arranged to logically combine outputs of the first comparator and the second comparator and produce a logic signal; and

a second OR gate arranged to logically combine the logic signals from the first OR gate and the AND gate and generate an indication for transmission of the link packet for the virtual lane (VL) on the given port.

3. The Link Packet Scheduler as claimed in claim 2, wherein said Link Packet Watchdog Timer is utilized to ensure that at minimum link packets will be independently scheduled at least once every 65,536 symbol times in accordance with the InfiniBand™ specification.

4. The Link Packet Scheduler as claimed in claim 1, wherein each of said credits is defined to be 64-bytes of available receive buffer space.

5. The Link Packet Scheduler as claimed in claim 1, wherein said link packet contains at least a Flow Control Total Block Sent (FCTBS) field which is generated by the transmitter logic and is set to the total number of blocks (64 bytes) transmitted since link initialization, and a Flow Control Credit Limit (FCCL) field which is generated by the receiver logic and is used to grant transmission credits to the remote transmitter.

6. The Link Packet Scheduler as claimed in claim 5, wherein said Flow Control Total Block Sent (FCTBS) field and said Flow Control Credit Limit (FCCL) field are used to guarantee that data is never lost due to lack of receive buffer at each end of a physical link.

7. The Link Packet Scheduler as claimed in claim 1, wherein said programmable utilization threshold is set such that the receive buffer has multiple data packets pending processing.

8. A data network, comprising:

a host system having a host-fabric adapter;

at least one remote system;

a switch fabric which interconnects said host system via said host-fabric adapter to said remote system along different physical links for data communications; and

one or more communication ports provided in said host-fabric adapter of said host system each port including a set of transmit and receive buffers capable of sending and receiving data packets concurrently via respective transmitter and receiver at an end of a physical link, via said switched fabric, and a flow control mechanism utilized to prevent loss of data due to receive buffer overflow at the end of said physical link.

9. The data network as claimed in claim 8, wherein said flow control mechanism is configured to support flow control through multiple virtual lanes VLs on the given port and to perform:

determining when a Link State Machine transitions into one of a ***LinkInitialize*** state, a ***LinkArm*** state and a ***LinkActive*** state, or when a configuration strap for enabling ***loopback*** operation changes states;

if the Link State Machine transitions into one of the ***LinkInitialize*** state, the ***LinkArm*** state and a ***LinkActive*** state, or when the configuration strap for enabling ***loopback*** operation changes states, scheduling a link packet transmission for all supported virtual lanes VLs on a given port;

if the Link State Machine does not transition into one of the ***LinkInitialize*** state, the ***LinkArm*** state and a ***LinkActive*** state, or when the configuration strap for enabling ***loopback*** operation does not change states, determining whether a Link Packet Watchdog Timer per virtual lane VL has expired;

if the per VL Link Packet Watchdog Timer has expired, scheduling a link packet transmission for that virtual lane VL;

if the per VL Link Packet Watchdog Timer has not expired, determining whether a receive buffer utilization indicating a data storage level of the receive buffer exceeds a programmable utilization threshold;

if the receive buffer utilization exceeds the programmable utilization threshold, prohibiting the link packet transmission for that virtual lane VL;

if the receive buffer utilization does not exceed the programmable utilization threshold, determining whether free credits accumulated exceeds a programmable credit threshold;

if the free credits accumulated exceeds the programmable credit threshold, scheduling a link packet transmission for that virtual lane VL; and

if the free credits accumulated does not exceed the programmable credit threshold, prohibiting the link packet transmission for that virtual lane VL.

10. The data network as claimed in claim 9, wherein each of said credits is 64-bytes from the receive buffer, and said credits are relinquished when data packets are removed from the receive buffer and that space is reclaimed as available for packet storage, or when link packets are received whose Flow Control Total Bytes Sent (FCTBS) fields differ from actual blocks received (ABR) at the given port.

11. The data network as claimed in claim 8, wherein said flow control mechanism contains a Link Packet Scheduler per virtual lane (VL) arranged to schedule a link packet transmission for that virtual lane VL.

12. The data network as claimed in claim 11, wherein said Link Packet Scheduler comprises:

a N-bit counter arranged to accumulate free credits relinquished, when a data packet is removed from a receive buffer and that buffer space is reclaimed as available for data packet

storage, or when a link packet is received whose Flow Control Total Bytes Sent (FCTBS) field differs from actual blocks received (ABR) at the given port;

a first comparator arranged to make comparison between accumulated free credits from the N-bit counter and a programmable credit threshold;

a second comparator arranged to make comparison between a current buffer receive utilization indicating a data storage level of the receive buffer and a programmable utilization threshold; and

a logic device arranged to track the current link state of the corresponding port, to monitor the amount of receive buffer resources from the first and second comparators and to schedule the transmission of a link packet, via a physical link.

13. The data network as claimed in claim 12, wherein said logic device comprises:

a first OR gate arranged to track whether a Link State Machine transitions into one of the ***LinkInitialize***, the ***LinkArm*** and the ***LinkActive*** state, whether a configuration strap for enabling ***loopback*** operation changes states, and whether a per-VL Link Packet Watchdog Timer expires at a predetermined symbol time, and to produce therefrom a logic signal;

an AND gate arranged to logically combine outputs of the first comparator and the second comparator and to produce a logic signal; and

a second OR gate arranged to logically combine the logic signals from the first OR gate and the AND gate and to produce an indication for transmission of the link packet for the virtual lane (VL) on the given port.

14. The data network as claimed in claim 12, wherein said Link Packet Watchdog Timer is utilized to ensure that at minimum link packets will be independently scheduled at least once every 65,536 symbol times in accordance with the InfiniBand™ specification.

15. The data network as claimed in claim 12, wherein each of said credits is defined to be 64-bytes of available receive buffer space.

16. The data network as claimed in claim 12, wherein said link packet contains at least a Flow Control Total Block Sent (FCTBS) field which is generated by the transmitter logic and is set to the total number of blocks (64 bytes) transmitted since link initialization, and a Flow Control Credit Limit (FCCL) field which is generated by the receiver logic and is used to grant transmission credits to the remote transmitter.

17. The data network as claimed in claim 16, wherein said Flow Control Total Block Sent (FCTBS) field and said Flow Control Credit Limit (FCCL) field are used to guarantee that data is never lost due to lack of receive buffer at each end of a physical link.

18. The data network as claimed in claim 12, wherein said programmable utilization threshold is set such that the receive buffer has multiple data packets pending processing.

19. A method of flow control of a link packet in a host-fabric adapter installed in a data network, comprising:

determining when a Link State Machine transitions into one of a ***LinkInitialize*** state, a ***LinkArm*** state and a ***LinkActive*** state, or when a configuration strap for enabling ***loopback*** operation changes state;

if the Link State Machine transitions into one of the ***LinkInitialize*** state, the ***LinkArm*** state and a ***LinkActive*** state, or when the configuration strap for enabling ***loopback*** operation changes states, scheduling transmission of a link packet for all supported virtual lanes VLs on a given port;

if the Link State Machine does not transition into one of the ***LinkInitialize*** state, the ***LinkArm*** state and a ***LinkActive*** state, or when the configuration strap for enabling ***loopback*** operation does not change states, determining whether a Link Packet Watchdog Timer per virtual lane VL has expired;

if the per VL Link Packet Watchdog Timer has expired, scheduling transmission of a link packet for that virtual lane VL;

if the per VL Link Packet Watchdog Timer has not expired, determining whether a receive buffer utilization indicating a data storage level of the receive buffer exceeds a programmable utilization threshold;

if the receive buffer utilization exceeds the programmable utilization threshold, prohibiting transmission of a link packet for that virtual lane VL;

if the receive buffer utilization does not exceed the programmable utilization threshold, determining whether free credits accumulated exceeds a programmable credit threshold;

if the free credits accumulated exceeds the programmable credit threshold, scheduling transmission of a link packet for that virtual lane VL; and

if the free credits accumulated does not exceed the programmable credit threshold, prohibiting transmission of a link packet for that virtual lane VL.

20. The method as claimed in claim 19, wherein each of said credits is 64-bytes from the receive buffer, and said credits are relinquished when data packets are removed from the receive buffer and that space is reclaimed as available for packet storage, or when link packets are received whose Flow Control Total Bytes Sent (FCTBS) fields differ from actual blocks received (ABR) at the given port.

21. The method as claimed in claim 19, wherein said Link Packet Watchdog Timer is utilized to ensure that at minimum link packets will be independently scheduled at least once every 65,536 symbol times in accordance with the InfiniBand™ specification.

22. The method as claimed in claim 19, wherein each of said credits is defined to be 64-bytes of available receive buffer space.

23. The method as claimed in claim 19, wherein said link packet contains at least a Flow Control Total Block Sent (FCTBS) field which is generated by the transmitter logic and is set to the total number of blocks (64 bytes) transmitted since link initialization, and a Flow Control Credit Limit (FCCL) field which is generated by the receiver logic and is used to grant transmission credits to the remote transmitter.

24. The method as claimed in claim 19, wherein said Flow Control Total Block Sent (FCTBS) field and said Flow Control Credit Limit (FCCL) field are used to guarantee that data is never lost due to lack of receive buffer at each end of a physical link.

25. The method as claimed in claim 19, wherein said programmable utilization threshold is set such that the receive buffer has multiple data packets pending processing.